

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

第2589299号

(45) 発行日 平成9年(1997)3月12日

(24) 登録日 平成8年(1996)12月5日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 7 1		G 1 0 L 3/00	5 7 1 C
	5 6 1			5 6 1 E
				5 6 1 G

発明の数1(全 9 頁)

(21) 出願番号	特願昭62-18078	(73) 特許権者	999999999 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22) 出願日	昭和62年(1987)1月28日	(72) 発明者	藤本 教幸 川崎市中原区上小田中1015番地 富士通株式会社内
(65) 公開番号	特開昭63-186298	(74) 代理人	弁理士 古谷 史旺
(43) 公開日	昭和63年(1988)8月1日	審査官	張谷 雅人
		(56) 参考文献	特開 昭62-262898 (J P, A) 特開 昭60-179836 (J P, A)

(54) 【発明の名称】 単語音声認識装置

1

(57) 【特許請求の範囲】

【請求項1】単語の使用頻度または重要度に応じて高順位の分類から低順位の分類までの複数の分類に分けられ、順位が高いほど単語音声のパターン数を少ないようにして単語音声のパターンの各々についての特徴を表すパラメータが登録されている複数の単語登録手段(111A, B, C, ……)と、  
入力単語音声のパターンについてその特徴を表す入力パラメータ(113)を得、複数の単語登録手段(111A, B, C, ……)のそれぞれが有する前記登録パラメータと照合し、距離若しくは類似度を求めて照合結果(115A, B, C, ……)として順次出力する照合手段(117)と、  
前記照合結果(115A, B, C, ……)を格納する格納手段(119)と、  
前記照合手段(117)から出力される前記照合結果(115

2

A, B, C, ……)を前記分類毎に前記照合結果(115A, B, C, ……)の距離または類似度に応じた順に、且つ前記照合結果(115A, B, C, ……)を各別にアクセス可能に格納手段(119)に格納させると共に、処理開始信号に応答して前記格納手段(119)に格納される最高順位の分類に含まれる照合結果の中から距離または類似度が第1位の認識候補の照合結果対応の単語を表す選択結果信号を出力し、該選択結果信号を認識結果の選択結果信号とすることができないとき、次候補要求信号(121)に応答して該次候補要求信号の送出時まで、前記格納手段(119)に格納されている照合結果のうちから既に認識候補として出力済みの照合結果を除いた中で距離の一番小さいもしくは類似度が最大の照合結果を選択し、該照合結果対応の単語を表す選択結果信号を認識結果として出力する選択制御手段(123)とを設けて構成したことを特

10

徴とする単語音声認識装置。

【請求項2】照合手段(117)は、複数の単語登録手段(111A,B,C,……)に対応した複数の照合回路部から成っており、該複数の照合回路部は入力パラメータ(113)を共通的に受け、各照合回路部は、複数の単語登録手段(111A,B,C,……)のうち対応する単語登録手段が有する単語の前記登録パラメータと照合して、前記照合結果を出力するように構成したことを特徴とする特許請求の範囲第1項記載の単語音声認識装置。

【請求項3】照合手段(117)は、1つの照合回路部で成り、複数の単語登録手段(111A,B,C,……)を順次切り換えて、各単語登録手段が有する前記登録パラメータと照合して、照合結果(115A,B,C,……)を順次出力するように構成したことを特徴とする特許請求の範囲第1項記載の単語音声認識装置。

【発明の詳細な説明】

【目次】

概要

産業上の利用分野

従来の技術

発明が解決しようとする問題点

問題点を解決するための手段

作用

実施例

I.実施例と第1図との対応関係

II.実施例の構成

III.実施例の動作

IV.実施例のまとめ

V.発明の変形態様

発明の効果

【概 要】

単語の使用頻度等に応じた順位で標準パターンを登録し、入力音声パターンと照合手段で照合する。その照合結果を前記順位毎に距離等に応じた順で格納手段に選択制御手段により格納する。格納手段の照合結果は、選択制御手段によって、最初は最高順位の照合結果のうちの距離等について第1位にある照合結果を出力し、認識結果とすることができないとき次候補要求信号に応答して該次候補要求信号送出時までの照合結果のうちの認識候補として出力済でない照合結果のうちの距離等から認識候補とし得る照合結果を認識結果として出力する。前記標準パターンの格納態様、照合結果の格納態様、照合結果の認識候補としての出力制御が相乗的に作用して認識率を向上させつつ、認識結果を短時間のうちにすることができる。

【産業上の利用分野】

本発明は、単語音声認識装置に関し、特に、人が発声する言葉を自動認識する技術である音声認識を適応し、登録されている音声パターンと照合して、発声された単語に関する情報を得るようにした単語音声認識装置に関

するものである。

【従来の技術】

従来から、このような音声認識に関する研究が盛んであり、また、それを応用した音声認識装置も開発、実用化されている。

このような音声認識装置の参考文献として、1983年11月7日発行の「日経エレクトロニクス」の第171頁～第208頁「連続発声した単語音声効率に認識する2段DPマッチング」が挙げられる。そこに紹介されている音声認識装置における音声認識処理としては、第3図に示すような流れとなっている。

図において、先ずマイクロホン451から入ってくる音声は、分析部453によって分析され、その音声パターンの特徴を表す認識パラメータが抽出される。

このシステムにあつては、特定話者用の単語音声認識装置であるとする、切換スイッチ455を「登録」の側に設定して、分析部453で抽出された音声パターンの特徴を表す認識パラメータを、その特定話者用に標準パターン部457に登録する。これにより、このシステムによって認識動作を行なう前に、その特定話者の各認識対象単語の分析結果が、標準パターンとして予め登録される。

実際に認識動作を行なうときには、切換スイッチ455を「認識」側に設定してある。各認識対象単語の標準パターン(標準パターン部457に登録済み)と、現入力音声パターン(分析部453から得られる)の両パラメータを比較して、最も近い(すなわち距離の小さい)認識対象単語を選択する。つまり、パターンマッチング処理を行なう。

ここで、パターンマッチング処理は、距離計算部459により、分析部453から得られる現入力音声パターンのパラメータと、既に標準パターン部457に登録されている各認識対象単語の標準パターンとの距離を演算する。また、最小値検出部461は、距離計算部459における計算結果に基づいて、最も距離の小さい標準パターン認識対象単語を抽出して、「認識結果」として出力する。

なお、パターンマッチング処理方法としては、距離計算手法の他に類似度計算手法も知られている。「距離の小さい」と、「類似度の大きい」ことは等価である。

【発明が解決しようとする問題点】

ところで、上述した従来方式にあつては、現入力音声パターンのパラメータを、標準パターン部457に予め登録してある認識対象単語の標準パターンと比較する際には、該標準パターン部457に登録してある認識対象単語の全てについて比較する。そのため、認識対象単語群の全てについて照合を行ない、1位、2位、3位、……を決定し、順番に「認識結果」として出力していた。

しかしながら、標準パターン部457に予め登録してある認識対象単語が少ないときには問題ないが、当該認識

10

20

30

40

50

対象単語が多いときには、それら認識対象単語の全てについて比較しているので、「認識結果」が得られるまでに多大の時間がかかる。そのため、認識動作における応答が遅くなってしまうという問題点があった。

通常、標準パターン部457については、その使用頻度を考慮しないで単語登録は行なわれている。

いま、多項目入力につき、それらについて認識動作を行なうものとする。

例えば、標準パターン部457に予め登録してある認識対象単語群での単語数が10000語であり、その内使用頻度の高い単語は1000語であるものとする。その場合、第3図に示すようなシステムでの認識性能は、使用頻度の高い1000語についての「認識率」が90パーセント、また、10000語の全てについての「認識率」は70パーセントであり、更に、1語当たりの「照合時間」は、0.5msであるものとする。

その場合の実効認識率は、70パーセントであり、また、応答時間は5秒（ $=0.5\text{ms} \times 10000\text{語}$ ）である。

このように、多項目入力として認識対象単語が多いときには、「認識結果」が得られるまでに多大の時間がかかってしまう。

本発明は、このような点にかんがみて創作されたものであり、実効認識率の向上を図ると共に、単語情報の照合に要する時間が短縮された単語音声認識装置を提供することを目的としている。

〔問題点を解決するための手段〕

第1図は、本発明の単語音声認識装置の原理ブロック図である。

図において、複数の単語登録手段111A,B,C,・・・は、単語の使用頻度または重要度に応じて高順位の分類から低順位の分類までの複数の分類に分けられ、順位が高いほど単語音声のパターン数を少ないようにして単語音声のパターンの各々についての特徴を表すパラメータを登録している。

照合手段117は、入力単語音声のパターンについてその特徴を表す入力パラメータ113を得、複数の単語登録手段111A,B,C,・・・のそれぞれが有する前記登録パラメータと照合し、距離もしくは類似度を求めて照合結果（115A,B,C,・・・）として順次出力する。

格納手段119は、前記照合結果（115A,B,C,・・・）を格納する。

選択制御手段123は、前記照合手段117から出力される前記照合結果115A,B,C,・・・を前記分類毎に前記照合結果115A,B,C,・・・の距離または類似度に応じた順に、且つ前記照合結果115A,B,C,・・・を各別にアクセス可能に格納手段119に格納させると共に、処理開始信号に応答して前記格納手段119に格納される最高順位の分類に含まれる照合結果の中から距離または類似度が第1位の認識候補の照合結果対応の単語を表す選択結果信号を出力し、該選択結果信号を認識結果の選択結果信号

とすることができないとき、次候補要求信号121に応答して該次候補要求信号の送出時までに前記格納手段119に格納されている照合結果のうちから既に認識候補として出力済の照合結果を除いた中で距離の一番小さいもしくは類似度が最大の照合結果を選択し、該照合結果対応の単語を表す選択結果信号を認識結果として出力する。

これらの構成要件により、本発明は構成されている。

〔作 用〕

入力パラメータ113が照合手段117に与えられると、該照合手段117は、複数の単語登録手段111A,B,C,・・・に単語の使用頻度もしくは重要度の順位順に従って予め格納されている各標準パラメータとの照合が照合手段117で行われる。その照合結果は、前記順位毎にその照合で得られる距離もしくは類似度順に選択制御手段123によって格納手段119に格納される。

このような格納が為されつつある間に、選択制御手段123によって格納手段119に格納された照合結果のうちの最高順位の照合結果に含まれるものであって、距離もしくは類似度が第1位の認識候補の照合結果に対応する単語を表す選択結果信号を出力する。

その出力を認識結果とすることができないとき、次候補要求信号に応答して前記格納手段119に格納済の照合結果のうちの認識候補として出力済の照合結果を除いた中で距離の一番小さいもしくは類似度が最大の照合結果を選択し、その単語を表す選択結果信号を認識結果として出力する。

前述のような標準パターンの格納態様、照合結果の格納態様、及び照合結果の出力制御が相乗的に作用して入力音声の認識率の向上を達成しつつ、認識結果を得るまでの時間を短縮させることができる。

〔実施例〕

以下、図面に基づいて本発明の実施例について詳細に説明する。

第2図は、本発明の一実施例における単語音声認識装置の構成を示す。

I. 実施例と第1図との対応関係

ここで、本発明の実施例と第1図との対応関係を示しておく。

単語登録手段111A,B,C,・・・は、第1パターン登録部211A,第2パターン登録部211Bに相当する。

入力パラメータ113は、区間検出力信号213における入力単語音声パターンの特徴を表す認識パラメータに相当する。

照合結果115A,B,C,・・・は、照合結果出力信号215に相当する。

照合手段117は、第1照合部217A,第2照合部217B,判定部218に相当する。

格納手段119は、照合結果格納部219に相当する。

次候補要求信号121は、キーボード241から与えられる次候補要求信号に相当する。

選択制御手段123は、判定部218、制御部223に相当する。

## II. 実施例の構成

以上のような対応関係があるものとして、以下本発明の実施例について説明する。

第2図に示す単語音声認識装置としては、特定話者用であるものとする。

マイクロホン231は、話者の音声を信号波形に変換するものであり、その波形信号は次のパラメータ抽出部233に供給されるようになってい

る。このパラメータ抽出部233は、それぞれ周波数帯域の異なるバンドパスフィルタを複数個設けておき、一定間隔でサンプリングするものである。

ここで、第1パターン登録部211Aおよび第2パターン登録部211Bとして設けられている2つの標準パターン登録部には、当該特定話者についての音声パターンの特徴を表す認識パラメータが、その特定話者用に登録されている。その登録方法としては、その特定話者がマイクロホン231に向かって通常の発声状態で発声する。その音声パターンの特徴を表す認識パラメータがパラメータ抽出部233によって抽出される。その抽出された音声パターンの特徴を表す認識パラメータが、当該特定話者用に第1パターン登録部211Aおよび第2パターン登録部211Bに登録される。かような登録動作により、この単語音声認識装置によって認識動作を行なう前に、その特定話者の各認識対象単語の分析結果が標準パターンとして予め登録される。

ここで、第1パターン登録部211Aおよび第2パターン登録部211Bの2つに登録単語を分ける基準は、当該特定話者に対する認識対象単語の使用頻度に従っている。例えば、全体として1000語を登録するものとして、その内の使用頻度の高い100語を第1パターン登録部211Aに登録し、これに対して使用頻度の高くない900語を第2パターン登録部211Bに登録する。

この単語音声認識装置としては、パラメータ抽出部233の後段に区間検出部235を設け、制御部223の制御の下に所定の区間について、パラメータ抽出部233で抽出されたパラメータを検出する。

この区間検出部235は、本来「音声」でない部分も音声波形に含まれているので、パワー等により、一定区間について区切って、「音声」の部分を取り出している。

その検出されたパラメータを表す区間検出力信号213が、第1照合部217Aおよび第2照合部217Bに共通に供給される。

この第1照合部217Aには、第1パターン登録部211Aに登録されている各認識対象単語の標準パターンが供給される。また、第2照合部217Bには、第2パターン登録部211Bに登録されている各認識対象単語の標準パターンが供給されるようになってい

る。第1照合部217Aおよび第2照合部217Bは共に制御部22

3の制御に基づいて、区間検出力信号213によって表される音声パターンの特徴を表す認識パラメータが、第1パターン登録部211Aに登録されている各認識対象単語の標準パラメータと、また、第2パターン登録部211Bに登録されている各認識対象単語の標準パターンとそれぞれ照合されて、単語毎に距離が求められ、その照合結果を表す照合出力信号214A、照合出力信号214Bが出力されて判定部218に供給される。

判定部218では、照合出力信号214A、照合出力信号214Bで表されるそれぞれの照合結果を受け取り、そのまま、照合結果出力信号215として、照合結果格納部219に順次格納されるようになってい

る。また、判定部218では、照合出力信号214A中の距離最小の単語を選択した後、出力制御信号216を、制御部223に供給すると同時に、第1位の認識結果として上記距離最小の単語を表す選択結果信号224が制御部223に供給される。

キーボード241は、この単語音声認識装置を操作するための多数のキーが具備しており、その中には、照合結果格納部219に格納された複数の認識対象単語を、任意に選択して制御部223が「認識結果」として、利用装置（図示せず）に与えられるようにするための次候補要求キー（図示せず）が含まれている。

第1位の認識結果が誤りであった場合には、使用者が、この次候補要求キーを押下することにより、制御部223から判定部218に次候補要求信号が送られる。判定部218では、照合結果格納部219から既に出力済みの単語を除いた中から距離最小の単語を選択し、選択結果信号224を制御部223に供給する。

## III. 実施例の動作

上述した構成による実施例の動作について、以下説明する。

この単語音声認識装置が対象としている特定話者が、マイクロホン231の前で、「認識動作」を行なうために、特定の単語を発声したものとする。

但し、「単語」は単音節のもの、また、それ以外のものも含むものとする。

マイクロホン231によって捕らえられた音声波形は、パラメータ抽出部233によって、音声パターンの特徴を表す認識パラメータが抽出される。その抽出された音声パターンの特徴を表す認識パラメータが区間検出部235に供給され、区間検出部235において、時間的にパワーの変化する特定の区間にてパラメータ検出され、その検出されたパラメータを表す区間検出力信号213が、第1照合部217Aおよび第2照合部217Bに共通に供給される。

制御部223から、第1照合部217Aおよび第2照合部217Bの照合動作を付勢するように制御信号が与えられる。

第1照合部217Aは、第1パターン登録部211Aに登録されている「高使用頻度の単語」音声と、区間検出力信号213として導入された検出単語音声パラメータと、それ

らの特徴を表すパラメータに基づいて比較する。第1パターン登録部211Aの登録単語は1000語と少ないので、全部の登録単語についての照合動作は速く、全ての照合に基づく照合出力信号214Aが第1照合部217Aから判定部218に供給される時間は短い。

また、第2照合部217Bも同様に、第2パターン登録部211Bに登録されている「低使用頻度の単語」の標準パラメータと、区間検出力信号213として導入された単語音声パラメータと照合する。ここで、第2パターン登録部211Bの登録単語は9000語と多いので、その照合動作は遅く、全てについての照合出力信号214Bが、第1照合部217Bから判定部218に供給される時間は長い。

制御部223によって制御される判定部218は、照合出力信号214Aおよび照合出力信号214Bを受け、照合結果出力信号215として、照合結果格納部219に与えられる。但し、「低使用頻度の単語」について格納の終了は遅い。

このとき、照合出力信号214Aに対応した「高使用頻度の単語」に対する「照合結果」は、その「距離」の小さい順に、第1位、第2位、第3位、……として、照合結果格納部219に格納される。

また、照合出力信号214Bに対応した「低使用頻度の単語」に対する「照合結果」も、その「距離」の小さい順に、第1位、第2位、第3位、……として格納される。但し、「高使用頻度の単語」に対する「照合結果」と、「低使用頻度の単語」に対する「照合結果」とは、それぞれの順に従っている。

判定部218からは、出力制御信号216が制御部223に与えられ、これによって、少なくとも最初の「照合結果」が判定部218において得られ、照合結果出力信号215として照合結果格納部219に格納されたことを通知することとなる。これを受けた制御部223は、まず、「高使用頻度の単語」に対する第1位の「照合結果」を照合結果格納部219から取り出すべく、判定部218に指令する。

判定部218は、「高使用頻度の単語」に対する第1位の「照合結果」を格納単語情報信号222として照合結果格納部219から求める。このようにして得た格納単語情報信号222に応じて選択結果信号224として制御部223に供給して、その次段に接続されるべき利用装置（図示せず）に「認識結果」として出力する。

仮に、この出力された第1位の「認識結果」が特定話者の意図した現発声単語でなければ、キーボード241に具わっている次候補キーを操作する。この次候補要求キーを操作するまでには、第2照合部217Bによっても照合動作が終了しているので、照合結果格納部219には、「高使用頻度の単語」のみならず、「低使用頻度の単語」の単語も照合結果格納部219に「照合結果」が格納されている。

従って、次候補キーが操作されれば、「高使用頻度の単語」に対する第1位の「照合結果」を除外し、その他の「高使用頻度の単語」および「低使用頻度の単語」の

中から、距離の小さい単語を判定部218は検索して格納単語情報信号222として得、選択結果信号224として制御部223に供給する。つまり、第2位の「認識結果」が第1位の「照合結果」を除いて求められる。

但し、第2位の「照合結果」が、特定話者の意図した現発声単語でなければ、再度次候補キーを操作することにより、第3位の「照合結果」を照合結果格納部219から取り出して、「認識結果」が利用装置に出力される。

以下、同様に、第4位、第5位、……と、キーボード241の次候補キーを操作することによって、任意に、照合結果格納部219に格納されている「認識結果」を取り出して利用装置に出力することができる。

このようにして、現に発声した特定話者の単語は、第1パターン登録部211Aに登録されていた「高使用頻度の単語」に対して正しい「認識結果」が得られる確率が高く且つその速度も速くなる。

つまり、現に発声した特定話者の単語は、第1パターン登録部211Aに登録されている「高使用頻度の単語」に対する照合結果、および、第2パターン登録部211Bに登録されていた「低使用頻度の単語」に対する照合結果と共に、「認識結果」として出力可能である。

従って、第1パターン登録部211Aに登録されている「高使用頻度の単語」は、1000語と少なく、その全単語の照合に要する時間は少ないので、この単語音声認識装置での特定話者に対する単語音声認識は素早くできることとなる。

#### IV. 実施例のまとめ

このように、予め利用頻度の相違に着目し、予め登録すべき単語をグループ分けし、標準パターンとして、第1パターン登録部211Aおよび第2パターン登録部211Bの双方に予め登録している。入力単語音声に基づく区間検出力信号213を照合する際、それが使用頻度の高いものであれば、直ぐに第1パターン登録部211Aの登録単語との照合結果が得られる。

つまり、略第1パターン登録部211Aに登録されている単語との照合に要する時間だけで、「認識結果」が得られるので、応答速度が速く且つ実効認識率が極めて高くなる。

ここで、従来との比較を示しておく。この単語音声認識装置にあっても、その個々の認識性能は同じと仮定する。つまり、使用頻度の高い1000語および10000語の全てについてのそれぞれの「認識率」は90パーセントおよび70パーセントであり、また、1語当たりの「照合時間」は、0.5msであるものとする。

この単語音声認識装置における実効認識率は、81パーセント（ $0.9 \times 0.9 = 0.81$ ）である。また、応答時間は0.5秒（ $0.5\text{ms} \times 1000\text{語}$ ）となる。但し、この時間は第1照合部217Aによって、第1パターン登録部211Aの登録単語と照合に要する処理時間であり、キーボード241の次候補キーを使用しなかった場合である。

このように、実効認識率の向上が図られ且つ単語情報の照合に要する時間が短縮されることが理解できるであろう。特に、入力項目が多くなればなる程この効果は顕著である。

#### V. 発明の変形態様

なお、上述した本発明の実施例にあっては、第 1 照合部 217A および第 2 照合部 217B の 2 つを単語照合手段として設けたが、これを 1 つの照合部としてもよい。その場合、制御部 223 の制御によって第 1 パターン登録部 211A および第 2 パターン登録部 211B をそれぞれ切り換えて、時間的にずれた形で、先ず第 1 パターン登録部 211A に登録されている使用頻度の高い各認識対象単語と照合する。続いて、第 2 パターン登録部 211B に登録されている使用頻度の低い各認識対象単語と照合するようにすればよい。「高使用頻度の単語」での「照合結果」が得られ、次候補要求キーを操作している間には、「低使用頻度の単語」の「照合結果」が得られているので、何ら不都合はない。

また、上述実施例にあっては、1 回の次候補キーの操作までに、「低使用頻度の単語」についての照合が完了しているものとしたが、必ずしも完了していなくてもよい。第 2 照合部 217B による照合結果を順次受け入れ、再度の次候補キー操作までに照合の終了している範囲内の照合結果に基づいて、距離の小さいものを順次「認識結果」とするようにすればよい。かような例は、「低使用頻度の単語」として定義した単語が極めて多い場合に起こり得る。

上述した本発明実施例にあっては、第 1 パターン登録部 211A および第 2 パターン登録部 211B に予め登録する各認識対象単語のグループ分けは、その使用頻度に基づいて行なうものとしたが、これに限られることはない。単語音声認識装置の利用の実情に合わせて、登録単語のグループ化を行なえばよい。このグループも 3 つ以上としてもよく、3 つ以上のパターン登録部を設けて登録し、その全てについて照合するようにしてもよい。

このグループ分けの基準として、「使用頻度」の他にも各種の基準が考えられる。例えば、「重要度」に基づき、音声認識装置の使用態様に応じてグループ分けしてもよい。

但し、例えば「緊急停止」等のような重要度の高い単語はその使用頻度は低い、「最重要度の単語」にグループ化しておく必要がある。

上述した実施例では距離計算手法を採用したが、本発明はこれに限られるものではなく、類似度の大きいものを求める類似度計算手法を採用することが可能であることは明らかである。

更に、「I. 実施例と第 1 図との対応関係」において、第 1 図と本発明との対応関係を説明しておいたが、これに限られることなく、各種の変形態様があることは当業者であれば容易に推考できるであろう。

#### 〔発明の効果〕

上述したように本発明によれば、音声の特徴を表す標準パラメータを使用頻度もしくは重要度の順位に従って登録し、その標準パラメータと入力音声パラメータとの照合結果を前記順位毎に、且つ照合結果の距離もしくは類似度の順に格納する。そして、このように格納される照合結果のうちの認識結果としての照合結果を出力するに当たって、先ず最高順位の照合結果に含まれ、距離もしくは類似度が第 1 位の照合結果対応の単語を表す選択結果信号を出力する。この出力を認識結果とすることができないとき次候補要求信号に応答して格納されている照合結果のうちの出力済でない照合結果であって、距離の一番小さいもしくは類似度が最大の照合結果を選択し、照合結果対応の単語を表す選択結果信号を認識結果として出力するようにしたので、前述の標準パラメータの格納態様、照合結果の格納態様、及び照合結果の出力制御が相乗的に作用して入力音声の認識率の向上を達成しつつ、認識結果を得るまでの時間を短縮させることができる。

#### 【図面の簡単な説明】

第 1 図は本発明の単語音声認識装置の原理ブロック図、第 2 図は本発明の一実施例による単語音声認識装置の構成ブロック図、第 3 図は従来から行なわれている音声認識の処理を示す構成図である。

図において、

111A, B, C, ……は単語登録手段、

113 は入力パラメータ、

115A, B, C, ……は照合結果、

117 は照合手段、

119 は格納手段、

121 は次候補要求信号、

123 は選択制御手段、

211A, B はパターン登録部、

213 は区間検出出力信号、

214A, B は照合出力信号、

215 は照合結果出力信号、

217A, B は照合部、

218 は判定部、

219 は照合結果格納部、

222 は格納単語情報信号、

223 は制御部、

224 は選択結果信号、

231 はマイクロホン、

233 はパラメータ抽出部、

235 は区間検出部、

241 はキーボード、

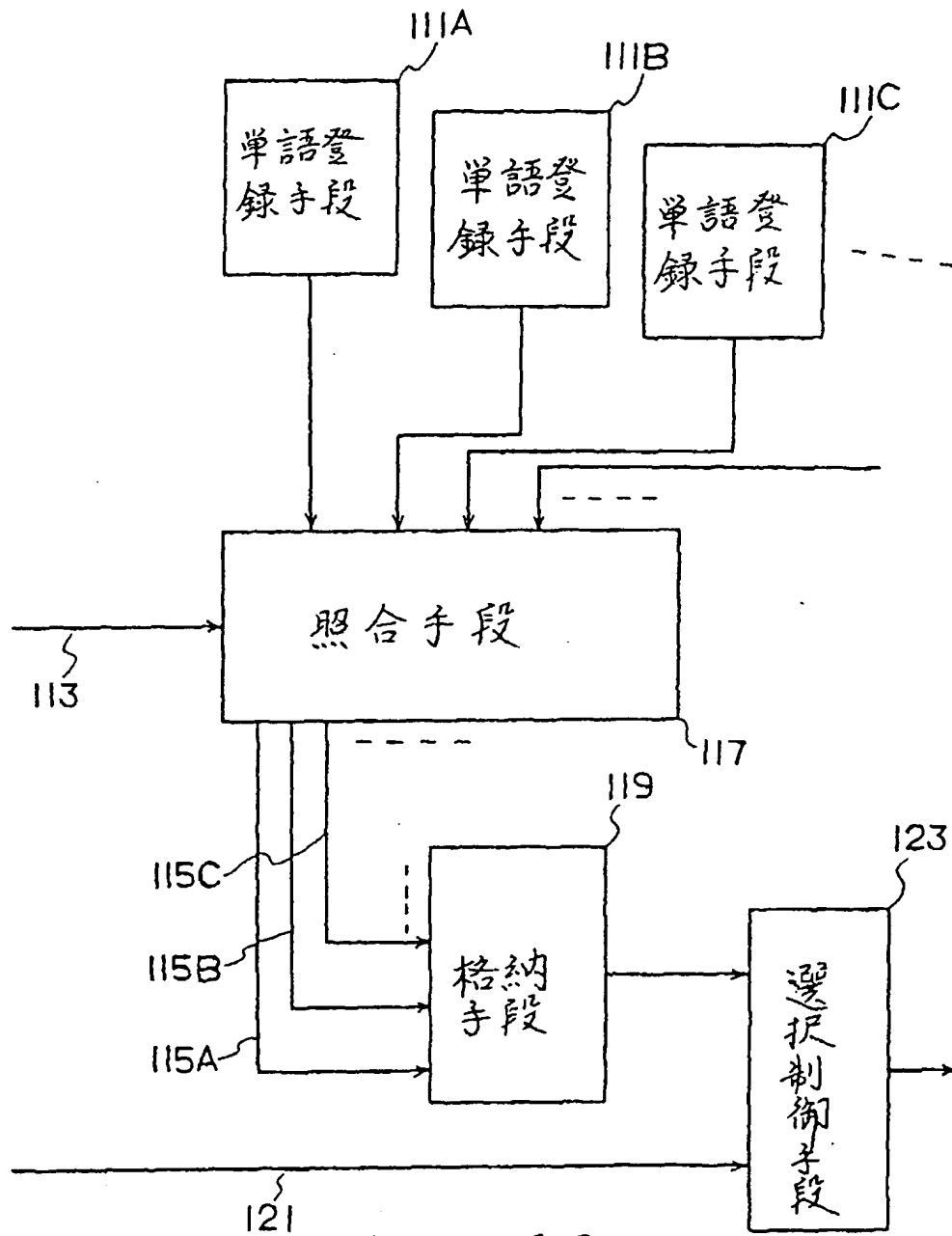
453 は分析部、

457 は標準パターン部、

459 は距離計算部、

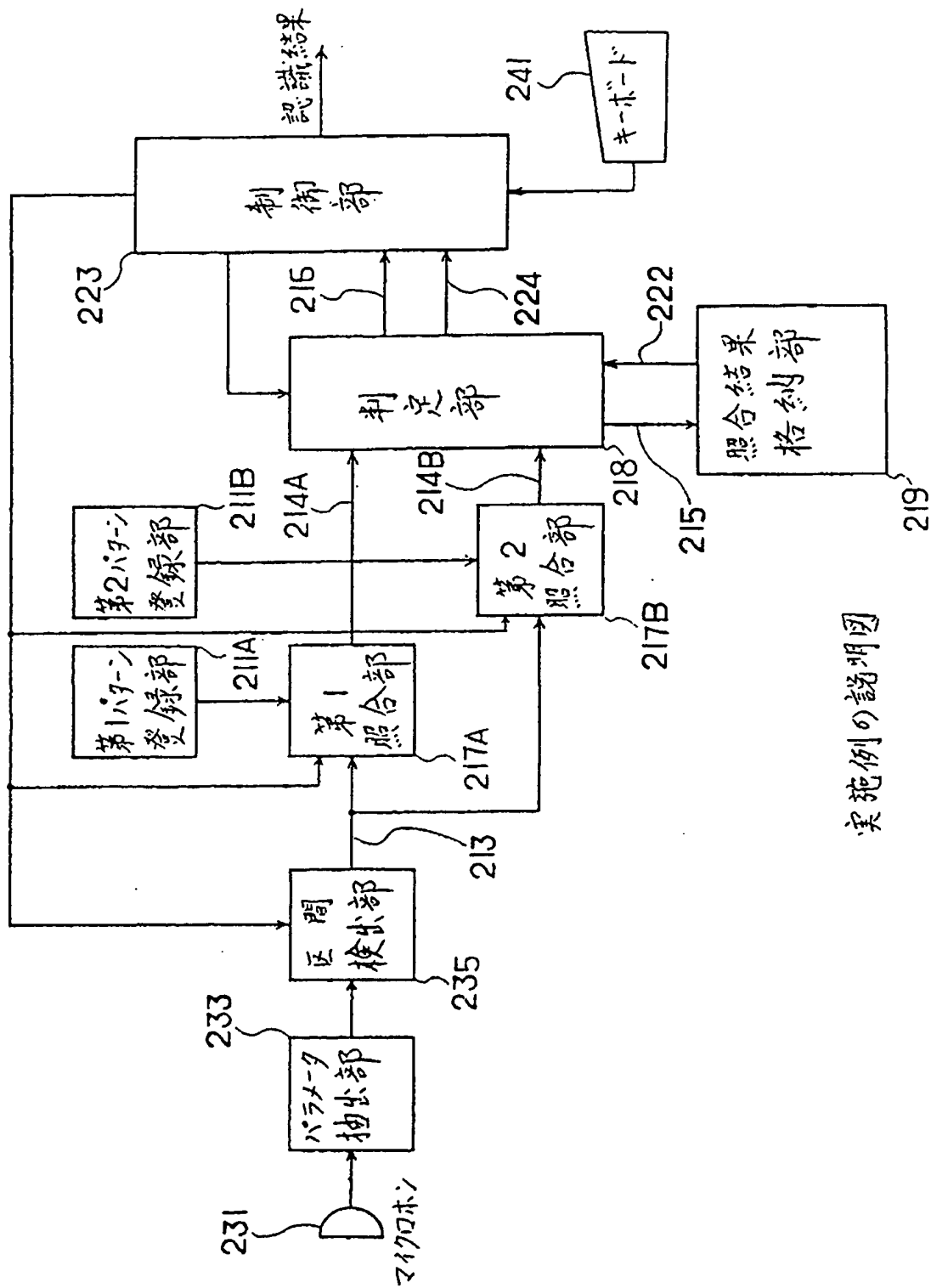
461は最小値検出部である。

【第1図】



本発明の原理ブロック図

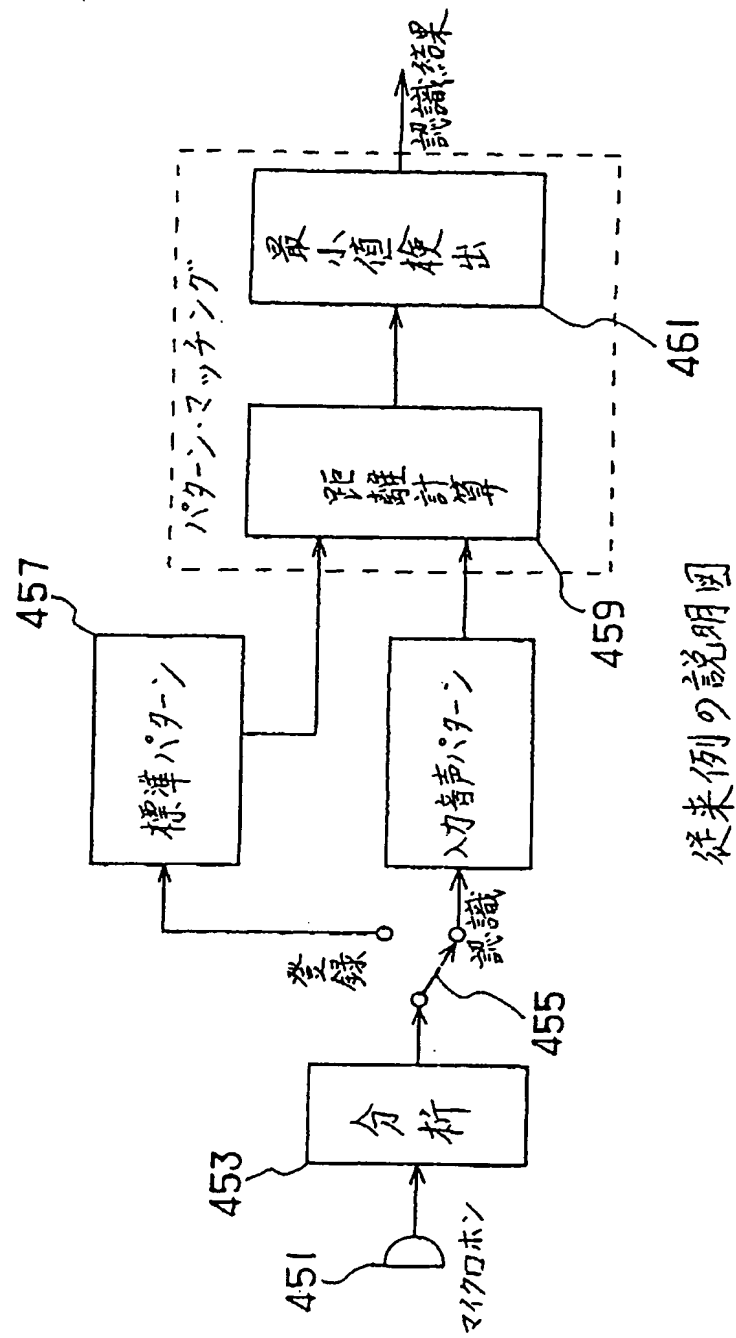
【第2図】



実施例の説明図



【第3図】



従来例の説明図